The Impact of Science Policy on the Rate and Direction of Cumulative Research: Frontier Tools and Applications

Jeff Furman, Boston University & NBER
Fiona Murray, MIT Sloan School
Scott Stern, Northwestern University & NBER

prepared for the Science of Science Policy
Workshop on the Science Of Science Measurement

version: Nov-20-2010 (preliminary)

Abstract

This paper outlines frontier tools for analyzing the impact of science policy on the rate and direction of cumulative research, describes several applications of those tools, and identifies series of research areas that could be usefully investigated with those tools. Specifically, we articulate an approach to evaluating science policy through the application of differences-in-differences techniques pioneered in the program evaluation literature to bibliometric data, which enable measurement of the creation and diffusion of scientific knowledge. We demonstrate how this approach has borne fruit in assessing the role of openness in scientific progress, the role of IP on cumulative knowledge, and the impact of restrictions on materials usage on research progress in human embryonic stem cell research. We conclude by offering ideas on three particular areas to which our methods could be usefully applied: (a) the study of national rules and regulations, (b) the study of agency- or university-level rules and norms, and (c) the study of community-level policy interventions.

I. Overview

Although the fact that knowledge is cumulative is recognized and understood to be central to economic growth, relatively little research has investigated the microeconomic and policy foundations of cumulative science and innovation (Mokyr 2002). For example, while "Open Science" is widely recognized as playing a central role in the production and diffusion of fundamental knowledge (Merton, 1973; Dasgupta and David, 1994; David 1998; Stephan, 1996), few formal analyses support our understanding of the impact of openness on the rate and direction of scientific progress.

In its effort to, "[support] research designed to advance the scientific basis of science and innovation policy," (NSF, 2010) we see the National Science Foundation's (NSF), Science of Science and Innovation Policy (SciSIP) program as a cornerstone effort to accelerate and expand the growth of knowledge around these topics. The program is an important and bold attempt to build a strong intellectual foundation science and technology policy making, grounded in empirical data, close observation of the practices and norms of the scientific community but also accounting for policy interventions, laws and rules that shape the institutional environment in which scientific research takes place. The explicit goals of the program are to fund research that "develops, improves and expands models, analytical tools, data and metrics that can be applied in the science policy decision making process" (NSF, 2010).

We conceptualize the SciSIP agenda as the systematic, evidence-based and causal analysis of the impact of policy interventions on the rate, direction on impact of scientific knowledge production. Overall, this agenda can enable policy-makers at the Federal and state level as well as others engaged in shaping the production and translation of scientific knowledge (including scientists themselves, universities, Foundations and scientific communities) to design more effective policies and practices that ensure that investments in science and innovation have rapid scientific, social and economic impact. Thus, the key "dependent variables" in SciSIP analyses include the rate of scientific knowledge production, the diversity of ideas, the quality of scientific knowledge (i.e. the extremes of the distribution), and the direction of knowledge production. This accords with the path breaking work of economists Kenneth Arrow and Richard Nelson who, in their 1962 volume the "Rate and Direction of Scientific Progress" argue for the importance of understanding "the supply of factors which are allocated to inventive effort, the output of inventive effort (inventions themselves) or the input-output relationship (the production function)." The key "independent variables" in SciSIP analysis include the set of policy levers – the intertwined set of institutional features - that might be used to shape the nature of knowledge producing activities. Broadly constituted, these levers include national rules and regulations, agency or university-specific rules and norms, community-based structures, rules or practices, and organizational interventions (at the lab level).

Leading scholars and policymakers have argued that the central challenge in evaluating science and technology policy is measuring the impact of public policies on research behavior, research outputs, and associated economic outcomes (Marburger, 2005; Jaffe, 2006). The particular research challenge is that of causal inference. Specifically, it is empirically difficult to separate the influence of a particular policy or institution on the growth or diffusion of knowledge from the underlying nature of the scientific knowledge that is being studied, even though the two are conceptually distinct. To put it more simply, the policy "whodunit" is that it is hard to know whether a particular policy or institutional environment has *caused* -- or was incidental to – the speeding up scientific progress in a particular area. Although social scientists and policy-makers hope to observe the *treatment effect* of a particular policy intervention (i.e., the causal impact of that policy), we may instead confuse this effect with a *selection effect*, by which certain types of knowledge end up associated with particular institutions or policy environments. For example, if we observe that U.S. universities have higher patent rates that E.U. universities, this fact may arise because the nature of the research done in U.S. universities differs from that of E.U. universities (the selection effect) rather than the fact that U.S. academics face different incentives to commercialize the research they conduct (the treatment effect). Without a parallel universe for policy experiments, when one observes the production of diffusion of a piece of knowledge within a given policy environment, one cannot directly observe the counterfactual production and diffusion of that knowledge had it been produced and diffused under alternative policy conditions.

An emerging area of work in SciSIP addresses these challenged by applying a citations-based 'natural experiments' approach to identifying treatment effects separately from selection effects. In the sections below, we provide an overview of the tools associated with this approach, noting its origins in literature in public finance and program evaluation, and summarize selected SciSIP research applying these tools.

II. A citation-based natural experiments approach to examining science policy

The substance of science policy takes many forms, from choices about the level of (and restrictions on) public funding, rules governing access to scientific research materials and data, and policies regarding intellectual property rights for discoveries resulting from (publicly funded) scientific process. Along each of these dimensions, science policy may influence both the overall productivity and the direction of the scientific research enterprise. In addition, science policy may have important *distributional* consequences: For example, while the establishment of “open access” research repositories may enhance accessibility and productivity for the *average* researcher, some researchers (e.g., those at leading institutions) may face a higher degree of scientific competition if key resources are made accessible to a wider set of researchers. Similarly, specific science policy interventions may enhance the impact of “important” discoveries while reducing attention to and/or diffusion of more minor findings.

The remainder of this section describes novel analytical tools that enable assessment of (a) the impact of science policy interventions on the rate and direction of scientific progress and (b) the distributional impact of science policy interventions.

II.1. The challenge of causal inference in program evaluation

The challenge of causal inference in program evaluation has been a topic of active study and substantial progress in economics over the past few decades (Imbens and Wooldridge, 2009). Early work addressing these in econometrics was motivated by the effort to understanding the impact of labor market programs, including worker training programs. Researchers in this area were confronted with the problem that the workers who selected into training programs were fundamentally different from (e.g., more motivated than or inherently more talented than) those that did not (Ashenfelter, 1978; Heckman and Robb, 1985). When participation into such programs is non-random, then a comparison of the outcomes associated with participation will conflate the underlying but difficult-to-observe characteristics of the individuals who participate in the programs with the causal effects of the programs themselves. Economists refer to these as problems of self-selection (or selection problems). More generally, economists refer to problems in which outcomes (such as the effects of labor market programs), measured as dependent variables, and drivers of those outcomes, measured as independent variables, as problems of endogeneity.

One method that economists have developed to mitigate these selection problems and obtain estimates of the causal impact of programs on outcomes is differences-in-differences estimation. This technique involves identifying a particular program (often referred to as a "treatment") and comparing differences in outcomes between those affected by the treatment (referred to as "the treated") and those not affected by the treatment (often referred to as the "controls") in time periods prior to and subsequent to the treatment. One important aspect of the technique is that the treatment itself should be exogenous – that is, in order to avoid the problems of selection, the arrival of the treatment should be random or, at the very least, not correlated with selection into the treated or control groups or anticipate changes in outcomes for these groups. Thus, this technique is often referred to as a 'natural experiments' approach, as empirical applications of the technique take advantage of unanticipated events that act as natural experiments on the populations affected by them.

II.2. The natural experiments approach in the context of SciSIP research

Natural experiments approaches applying differences-in-differences estimation have been particularly prevalent in labor economics, public finance, and development economics, each of which is a field in which assessing the causal impact of policies or programs is of central concern (Meyer, 1995; Bertrand, Duflo, and Mullainathan, 2004; Angrist and Pischke, 2008).

Increasingly, these techniques are being applied in SciSCIP research. Often, the application of these techniques involves the use of bibliometric data, using data on published research papers and the citations to such papers.

This approach is premised on several assumptions. The first assumption is that data on the production and citation of academic papers provides valuable (though imperfect) indicators of scientific progress. The seminal work of Merton (1973), Garfield (1950), and De Solla Price (1971) articulates the importance of priority and publication in the system of scientific rewards and noting the importance of publications and citations in tracking the rate and direction of scientific progress.[1] The technique relies upon the fact that (a) academic papers are produced at a specific and measurable point in time and (b) the use by follow-on researchers of the knowledge articulated in those papers takes place over time and in a way that can be measured as well. The approach interprets citations to academic papers as evidence of the use of prior knowledge by follow-on researchers, although we acknowledge that these are noisy measures.

The second assumption is that the degree to which future research "draws upon" (cites) a given article (and by whom and where and when) depends on institutional mechanisms, including intellectual property rights over the knowledge disclosed in the article, rules and institutional arrangements governing access to research resources, and national and local policies. Further, the opportunity to take advantage of a given research trajectory by researchers in any one location or institution depends on access to funding, materials, and support infrastructure to conduct that research in a timely manner (i.e., before others are able to exploit the opportunity). The impact of institutions and policy interventions on facilitating this process of step-by-step scientific discovery is a key challenge for science policy, and a central focus of science policy analysis (Aghion Dewatrapoint & Stein 2008; Mokyr 2002).

The third and final assumption is that science policy interventions that change the institutional environment for scientific research will be reflected in changes in the rate and direction of scientific progress, which, in turn, is captured in citation patterns. To ensure that researchers can recover a direct estimate of the impact of the institutional or policy change on the treated sample, they often employ a carefully-chosen control sample that, in theory, should not have been

[1] The field recognizes that bibliometric analysis provides only a noisy indicator of scientific progress (see, e.g., Garfield (1979), Lindsey (1989), and Schubert and Braun (1993)): For a number of reasons, small differences in the citation rate of a single paper (particularly early in its publication history) are of limited value in distinguishing the importance of research or its use by the research community. This research approach takes care to minimize the impact of these limitations by drawing comparisons among large samples of publications, comparing across control samples, and assessing the impact of policy changes by drawing comparisons within articles across time. Moreover, the approach should deal with these problems by averaging them out across the areas it studies, unless the problems are closely correlated with the specific policy or institutional changes we study.

affected by the policy under study. Under certain circumstances, then, the research is able to estimate the impact of the policy even in the absence of a control sample.

II.3 Citation analysis via differences-in-differences estimation

To apply this methodological approach, SciSIP researchers would ideally like to observe a given piece of knowledge in multiple institutional or policy environments and compare the generation or diffusion knowledge across those environments. In practice, it is not possible to assign 'units of knowledge' to separate environments, so the analytical framework relies on the fact that institutional changes or policy interventions may induce changes in the production of scientific articles or changes in citation behavior relative to baseline levels.

It is essential that such environmental changes do not impact the original knowledge itself but do impact the incentives and opportunities for follow-on researchers to exploit that knowledge after the environmental change. Thus, the approach investigates the extent to which policy interventions and changes in institutional environments induce changes in the production of or citations to scientific articles in the pre- and post- policy (treatment) period, relative to a set of control articles which are not impacted by the policy treatment.

To do this, this approach employs an estimation technique that identifies the average differences in citations across the affected groups ("treated groups") and unaffected groups ("control groups"). It then estimate the year-to-year change in citations resulting from the change in the institutional or policy environment.[2] The analysis takes care to control for a number of factors, including the average importance of each article, the calendar year in which articles are cited, and the age of the article, along with other relevant factors. Thus, one tests for the impact of policy interventions by calculating how the citation rate for a scientific publication *changes* following such interventions, accounting for fixed differences in the citation rate across articles and relative to the non-parametric trend in citation rates for articles with similar characteristics.

[2] Specifically, this baseline estimator is:

$$CITES_{i,j,pubyear(j),t} = f(\varepsilon_{i,j,t};\gamma_i + \beta_t + \delta_{t-pubyear} + \psi POST-TREATMENT_{i,t}),$$

where (γ_i) is a fixed effect for each article, β_t is a year effect, $\delta_{t\text{ - pubyear}}$ captures the age of the article, and POST-TREATMENT is a dummy variable equal to one only for years after the knowledge linked to the article is affected by the institutional or policy change. The coefficient on POST-TREATMENT (ψ) indicates the marginal impact of the intervention on the set of treated articles.

The core estimation technique can be modified to address the possibility that citation rates may change prior to or after the policy shift. This is important for understanding the dynamic consequences induced by the policy intervention – for example, whether the impact of the policy intervention occurs as a one-time change in the levels or diffusion of knowledge, whether it declines or returns to baseline over time, or whether the policy intervention induces continuously growing effects. If the changes in citation patterns occur prior to the policy shift, this could indicate that the observed changes are not the *result* of the policy, but actually pre-date the policy and, therefore, may be unassociated with the policy intervention.

Thus far, the analysis describes the ways in which policy shifts affect the overall levels of knowledge generated or diffused. By taking advantage of bibliometric details, it is possible to examine the impact of policy interventions not just on the overall amount of research, but on *specific populations* of researchers, such as researchers from particular geographic regions, research engaged in various types of collaborations, researchers in different institution types (academic, government, private), researchers from elite or non-elite universities, or research in basic or applied journals. To estimate the impact of policy interventions on each of these subpopulations, one can aggregate these individual citations into counts of the number of citations received by a given article in a given year by a given subpopulation of citers.[3] Equally, one could group citations to particular article types, thus tracking the impact of policy interventions on specific subpopulations of research by specific subpopulations of researchers, thus enabling an extremely nuanced assessment of the impact of a particular policy shift on the scientific community.

III. Applications

In this section, we introduce a number of contexts in which the research approach we outline above has been employed to evaluate selected science policy interventions. One of the overarching findings of this research stream is that the active management of the scientific

[3] In this formulation, we estimate the equation:

(2) $$CITES_{i,l,t} = f(\varepsilon_{i,j,t}; \gamma_i + \lambda_l + \beta_t + \delta_{t-pubyear} + \sum_{l=1,...L} \psi_l \iota_l POST-TREATMENT_{i,t})$$

In this equation, ψ_l represents the average impact of the treatment on sub-population *l*, conditional on a fixed effect for each article, age and citation-year.

environment can have an impact on the rate and direction of cumulative research, although strategic behavior by scientists, academic institutions, and other actors can limit the effectiveness of federal science policy interventions.

III.1 The Impact of Openness on the rate and direction of scientific progress

As we noted in the introduction to this paper, "Open Science" is widely recognized as playing a central role in the production and diffusion of fundamental knowledge; however, few quantitative analyses document impact of openness on the rate and direction of scientific progress. In "Climbing atop the shoulders of giants: The impact of institutions on cumulative research," Furman and Stern (2010), we evaluate the impact of moving biomaterials from peer-to-peer collections, to centralized, public Biological Resource Centers, such as the American Type Culture Collection (ATCC), on the rate of cumulative research exploiting those biomaterials. We take advantage of a number of unusual events in which a large number of materials, which has previously circulated in the academic community, were simultaneously and unexpectedly shifted from smaller private collections into the ATCC, the United States' largest public culture collection. We find a dramatic increase in research impact when biomaterials are made accessible on a certified, standardized, and low-cost basis through ATCC. We also find evidence that the movement of research materials into ATCC has a 'democratizing' impact – e.g., that research initially conducted at lower status institutions, lower-impact journals, and geographic regions outside United States experienced a greater increase in follow-on work than research at higher status institutions, higher-impact journals, and US-based researchers (Furman and Murray, 2010).

We further investigate the impact of openness on the rate and nature of scientific progress in, "Of Mice and Academics: The Role of Openness in Science," (Murray et al. 2008). This paper explores what types of researchers and research projects are most likely affected by changes in openness in mouse genetics research. Specifically, we examine the impact of the DuPont-NIH Memorandum of Understanding in 1999/2000 reducing enforcement of the Oncomouse (and Cre-lox) patents. Using a sample of engineered mice that are linked to specific scientific papers (some affected by the NIH agreements and some not), we implement a differences-in-differences estimator to evaluate how the level and type of follow-on research using these mice changes after the NIH-induced increase in openness. We find a significant increase in the level of follow-on research. Moreover, this increase is driven by a substantial increase in the rate of exploration of *more diverse* research paths. Specifically, the bulk of the increase in citations arises from articles that are published by “new” researchers, institutions, and journals. The boost in incremental citations to a given mouse-article is greater among researchers working at institutions that had not cited that mouse-article prior to the NIH agreement, in journals that had not previously cited the article, and in among articles with previously unused “keywords” describing the underlying research contributions of the citing articles. Taken together, these findings suggest that openness increases the diversity of experimentation that follows from a single idea.

III.2 The Impact of IP on Scientific Progress

A central question in science and innovation policy regards the impact of intellectual property choices on the rate at which future researchers build on ideas protected by patents or other intellectual property rights. In a set of papers, we apply the methodology outlined above to analyze the impact of natural experiments in IP protection. In “Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge?” (Murray & Stern 2007) examines the impact on knowledge flows associated with ideas described in papers published in the journal *Nature Biotechnology* to which patent protection is subsequently extended. The paper demonstrates that over 50% of publications in *Nature Biotechnology* are associated with such “patent-paper pairs.” Patents are granted on the ideas described in the original papers approximately three to four years after publication. Thus, knowledge associated with those papers is exposed to two distinct environments: the pre-grant period where informal norms hold and the post-grant period when IP rights can be enforced. Using a differences-indifferences estimator the paper finds a 10-20% decline in the citation rate to the paired publication after patent grant, suggesting that patenting limits but does not eliminate follow-on research.

Building on this initial paper, “Learning to Live with Patents: Assessing the Impact of Legal Institutional Change on the Life Science Community” (Murray & Stern 2008), suggests that academics at high-status (or elite) academic institutions are better able to adapt to the imposition of IP rights than counterparts at lower-status institutions, a finding which highlighting the importance of the social organization of science in shaping the response to IP.

Heidi Williams' work provides an additional example of the natural experiments approach to analyzing the impact of IP on knowledge generation and diffusion (Williams, 2010). Williams identifies sets of genes sequenced by the public research effort (the Human Genome Project, HGP) and the for-profit research effort (Celera) that were, in principle, equally-likely targets for scientific research and clinical products. Her analysis compares the extent of cumulative advances on research characterizing HGP- and Celera-sequenced genes, using academic papers, patents, and commercial outputs, such as diagnostic tests as measures of follow-on progress. She finds documents a substantial reduction in cumulative research in genes initially characterized by Celera, whose use was thus governed by Celera's intellectual property terms. Indeed, she documents a greater than 30% reduction in the probability that a gene-specific diagnostic test has been developed for genes that were initially sequenced by Celera.

III.3 The Impact of Restrictions on Research Materials – the case of Federal Stem Cell Policy

Traditional studies of scientific research highlight the influence of different organizational and institutional conditions imposed by funders (corporate or government) on researcher productivity. In "Growing Stem Cells: The Impact of U.S. Policy on the Organization of Scientific Research," (Furman, Murray, and Stern, 2010a), we argue that this approach should acknowledge scientists' robust preferences to freely select their research direction and shape their organizational context. As a window into the importance of such strategic choices, we examine the scientific community's response to the U.S. Administration human embryonic stem cell (hESC) research policy. This policy constitutes an experiment that allows us to investigate the way in which restrictions on U.S. scientists' freedom affected their participation in and organizational choices regarding research. Our estimates suggest that in the aftermath of the 2001 policy, U.S. production of hESC research lagged 35-40 percent behind anticipated levels. However, the decline was largely concentrated in the years 2001-2003 with the impact later ameliorated, in particular among researchers at elite U.S. institutions and among U.S. researchers who collaborated with international partners. These results suggest that scientists do indeed engage in strategic action - both in terms of exit but also action - when faced with limits to their freedom. For scholars focusing on the role of freedom in the scientific community, our results can be interpreted as affirmative evidence. Moreover, they speak to the challenges associated with using intervention as a tool to shape the rate, direction and organization of science.

IV. Implications from extant research

In this paper, we describe novel tools for quantitative analysis of the impact of science policy interventions on the process of cumulative scientific discovery. This analysis extends prior research by exploiting the recent availability of detailed citation data with frontier methods from the program evaluation literature. The approach moves beyond traditional cross-sectional comparisons of citations associated with knowledge in different institutional or policy environments; instead, we focus on "natural experiments" where the conditions governing access, diffusion or follow-on research funding associated with a given piece of knowledge are changing over time. This approach allows researchers to disentangle the role of institutions and policy in shaping scientific progress from the intrinsic variation in scientific importance across discoveries. In addition, the tools also enable the evaluation of the *distributional* consequences of policy initiatives.

We propose that there are at least two implications of research in this area so far for agencies that fund scientific research and are can affect science policy. The first is methodological: The results suggest that the citations-driven, natural experiments approach to science policy can be useful for studying the generation and diffusion of cumulative research. The careful design of scientific funding and research and researcher tracking can facilitate the science of science policy by enabling additional research applying the tools we outline above. For example, one of the great difficulties in evaluating the impact of policy interventions on individual scientists is the ability to track individual scientists. Indeed, the examples of differences-in-differences analyses of science policy interventions generally involve field-level data, as a consequence of the difficulty of tracking individual scientists and matching them to specific research output. Azoulay, Graff-Zivin, and Manso, 2010, and Azoulay, Graff-Zivin, and Sampat, 2010) are exceptions, notable for the lengths to which the researchers must go in order to ensure that they can track individual biomedical researchers. As another example, it is often difficult to evaluate the impact of funding policy decisions that involve, say, choices to fund particular projects rather than others, as a consequence of the difficulty in identifying projects that did not receive funding. While it is often not practical to make such data available, it could prove useful to policy evaluations to design data structures and funding awards in ways that facilitate analyses of the type we describe in this paper.

The second implication is more broad-reaching. The results suggest that institutional conditions and policies can, indeed, have a substantial impact on the both the rate and the nature of knowledge accumulation and diffusion. While policy debates often focus on the extent of investment in science and innovation, our research suggests that nuanced policies that encourage openness, ensure that research materials and tools are made available for follow-on research, and provide incentives for translational impact can have a substantial impact on knowledge accumulation. The findings from this research area suggest that government agencies, academic institutions, and individual scientists can develop institutions and practices that maximize long-term research productivity and enhance the likelihood of effective clinical translation.

V. Extensions and Research Agenda

The SciSIP agenda presents far reaching research opportunities for scholars whose goal is to contribute to the social sciences, to enhance understanding of the role science and innovation in the economy and, to have broader impact on society via public policy evaluation. The promising results of the findings described above suggest that the citations-based natural experiments approach is a promising methodology for assessing the impact of science policy. A number of important gaps in the current state of knowledge remain. We describe below a number of areas of research that could be usefully addressed with the approach we articulate above. We organize these according to the level of analysis at which the policy interventions are taking place: national rules and regulations, agency-level interventions, community norms and practices and organizational actions.

V.1. National rules and regulations

This area of potential future research would examine the effectiveness of national rules and regulations on the rate and direction of scientific progress. For example, a substantial amount of has examined on the effect of the Bayh-Dole Act on universities and university researchers (see e.g., Mowery et al., 2001, and Owen-Smith and Powell 2003). Gaps at this level of analysis remain with regards to the role of international rules and regulations on science in the United States, and the ability of U.S. researchers to remain highly competitive and at the knowledge frontier in the light of growing global spending on scientific research. In addition, it would be valuable to understand how the particular structure and incentives of university systems in different countries influence the impact of government policies such as those related to intellectual property rights.

Broad questions also remain regarding the overall impact of national funding for progress in particular scientific areas. For example, little is yet known about the rate of return, in terms of diagnostics and treatments (or even papers and patents), on investments in the NIH. Such work is difficult to do at the most aggregate level, i.e., for the economy overall. However, by identifying particular areas affected by particular policies may offer a promising route for evaluating specific country-level funding decisions, which could form the foundation for understanding in this area.

V.2. Agency or university-level rules and norms

Funding agencies, especially the Federal government, have a variety of opportunities to shape the rate and direction of scientific progress. Recent work funded by SciSIP has made significant progress along these two dimensions but gaps remain. In particular, the influence of non-Federal funding sources, especially corporate funding and foundation funding, is poorly documented and understood.

As they select among research projects and shape the expectations and controls they place on researchers funding agencies have a variety of opportunities to influence knowledge production. This has often been thought of as a black-box with the scientific community utilizing the peer review system as the best mechanism to self-regulate and shape direction. Recent analysis by Azoulay and colleagues (2010) applying techniques similar to those we outline above demonstrates that scientists are much more likely to produce innovative breakthrough science when they are subject to long-term grants that allow them exceptional freedom in the lab rather than grants that focus on more near-term outcomes. This study raises the question of how researchers are encouraged to move into new and emerging research areas, and how to encourage ideas at the high-quality, high-risk tail of the distribution.

It would be beneficial to encourage more research to understand the impact of funding choices and funding incentives on the type of research outcomes. This agenda could also benefit from the analysis of scientists outside the U.S. in settings where different types of incentive systems exist. In line with recent interest in challenges (prizes) as an alternative incentive mechanism, we should also extend this analysis to include other funding mechanisms or reward systems.

In addition to affecting science outcomes by affecting the extent of funding available for various research areas, funding agencies have an opportunity to shape the rate and effectiveness with which knowledge is generated through choices about the disclosure of knowledge and sharing of research materials and outputs. Among the most important and controversial rules shaping such impact of scientific research are the rules around intellectual property rights. This has been the topic of vigorous debate particularly with regards to the increasing levels of patenting within the scientific community. This is the research arena in which SciSIP researchers have made one of the greatest contributions, with their research informing policy discussions at the National Academies of Science, within the National Institutes of Health (NIH) and elsewhere. In particular, research has explored the impact of patenting on the rate at which that research is diffused within the scientific community and on the rate at which commercial or socially-beneficial products are developed (Murray and Stern 2007; Huang and Murray 2009; Walsh et al. 2003, 2005). Extensive research documents the impact of IP, licensing and material sharing practices on scientists, but gaps in our knowledge exist with regards to the impact of these policies on both scientific knowledge production and economic impact (few studies examine both with Williams (2010) a notable exception). However, we also have a less systematic understanding of how to design the "intellectual commons" in an efficient and effective manner so as to promote and rapid follow-on research and commercialization (Furman, Murray, and Stern, 2010b). There is also a significant opportunity to extend these studies beyond the study of life scientists to explore differences across research communities in a range of disciplines such as chemistry, computer science, materials science etc.

V.3. Community-level activities

The policies and practices that emerge from the scientific community also play a critical role in scientific progress and impact. There is definitive evidence that investments in community-based infrastructure such as materials repositories and data repositories have a significant positive impact on the rate of scientific progress by enabling access, certification and sharing (Furman and Stern 2010). As well, analysis of the self-governance of scientific communities through the system of retractions has also pointed out the role of the community as a crucial analytic lens (Furman and Murray 2009). In another stream of research grounded in organizational theory and sociology, scholars have examined whether and how different community structures emerge in order to undertake the complex task of horizontal collaboration (e.g. Powell et al. 2004) and collective work (Ferraro and O'Mahony forthcoming).

At this level of analysis, critical questions remain unanswered: How do science policies affect the formation and evolution of scientific communities? How do they coalesce around new research areas and what role might policy-makers play in such community formation? For example, do mechanisms such as those used in DARPA enable community building and how does this shape the long run effectiveness of scientific communities? (See Fuchs, 2010 for an example of work in this area.)

References

Aghion, P., M. Dewatrapoint, J. Stein. (2008). "Academic Freedom, Private-Sector Focus and the Process of Innovation," *Rand Journal of Economics*, 617-635.

Angrist J. and J.S. Pischke (2008) *Mostly Harmless Econometrics*, Princeton, NJ: Princeton University Press.

Ashenfelter, O. (1978) "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 6(1): 47–57.

Azoulay, P. G. Manso, and J. Graff Zivin (2010) "Incentives and Creativity: Evidence from the Academic Life Sciences," MIT Sloan Working Paper.

Bertrand M., Duflo E. and Mullainathan S. (2004) "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 119(1) 249-275.

Dasgupta, P. and P. David (1994) "Towards a New Economics of Science," *Research Policy* (23), 487-521.

David , P.A. (1998). "Common Agency Contracting and the Emergence of "Open Science" Institutions," *American Economic Review – Papers and Proceedings*, 88(2): 15-21.

De Solla Price, D. (1970), "Citation Measures of Hard Science, Soft Science, Technology, and Nonscience", in , C. Nelson and D. Pollock, eds. *Communication among Scientists and Engineers*, (Cambridge: Heath Lexington Books).

Ferraro F. and S. O'Mahony, (2010) "Managing the Boundary of an Open Project," in *Market Emergence and Transformation*, W. Powell and J. Padgett, Eds., forthcoming.

Fuchs, E. (2010) "Rethinking the Role of the State in Technology Development: DARPA and the Case for Embedded Network Governance," *Research Policy*, 39(9), 1133-1147

Furman, J. and S. Stern. (2010) "Climbing Atop the Shoulders of Giants: The Impact of Institutions on Cumulative Knowledge Production," *American Economic Review*, forthcoming.

Furman, J., F. Murray, and S. Stern. (2010a) "Growing Stem Cells: The Impact of U.S. Policy on the Organization of Scientific Research," working paper.

Furman, J., F. Murray, and S. Stern. (2010b) "How to govern the 'intellectual commons'?," working paper.

Garfield E. (1955) "Citation Indexes for Science," *Science* 122: 108-11.

Griliches, Z. (1990) "Patent Statistics as Economic Indicators: A Survey," *Journal of Economic Literature*, 92: 630-653.

Griliches, Z. (1998) *R&D and Productivity: The Econometric Evidence*, University of Chicago Press.

Huang, K., F. Murray (2009) "Does Patent Strategy Shape the Long-Run Supply of Public Knowledge: Evidence from Human Genetics," *Academy of Management Journal*, 52(6), 1193 – 1221.

Heckman, J. J., and R. Robb (1985) "Alternative Methods for Evaluating the Impact of Interventions." in *Longitudinal Analysis of Labor Market Data*, ed. J. J. Heckman and B.Singer, 156-245. Cambridge, UK: Cambridge University Press.

Heller, M. A., R. Eisenberg (1998) "Can Patents Deter Innovation? The Anticommons in Biomedical Research," *Science* 280(5364): 698-701.

Imbens, G. W. and J. M. Wooldridge (2009) "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, 47(1), 5–86.

Jaffe, A., M. Trajtenberg, and R. Henderson (1993), Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations, *The Quarterly Journal of Economics*, pp. 577-598.

Jaffe, A. B. and J. Lerner (2004) *Innovation and Its Discontents: How our broken patent system is endangering innovation and progress, and what to do about it.* Princeton: Princeton University Press.

Jaffe, A. (2006) “The 'Science of Science Policy': Reflections on the Important Questions and the Challenges They Present," Keynote Address at the NSF Workshop on Advancing Measures of Innovation: Knowledge Flows, Business Metrics, and Measurement Strategies.

Lindsey, D. (1989). Using citation counts as a measure of quality in science: Measuring what's measurable rather than what's valid. *Scientometrics* , 15, 189-203.

Marburger, J. (2005) “Wanted: Better benchmarks,” *Science,* 308(5725): 1087.

Merton, R. (1957/1973) *The Sociology of Science: Theoretical and Empirical Investigations*. Edited by Norman Storer. Chicago: University of Chicago Press.

Merton, R. (1973) *The Sociology of Science: Theoretical and Empirical Investigations*. Edited by Norman Storer. Chicago: University of Chicago Press.

Meyer, B. (1995) “Natural and Quasi-Natural Experiments in Economics,” *Journal of Business and Economic Statistics*, 13, 151-162.

Mokyr, J. (2002) *The Gifts of Athena: Historical Origins of the Knowledge Economy.* Princeton University Press, Princeton.

Mowery, D. C., R. R. Nelson, B. N. Sampat and A. A. Ziedonis (2001) “The growth of patenting and licensing by U.S. universities: an assessment of the effects of the Bayh–Dole act of 1980,” *Research Policy,* 30 (1), 99-119

Mukherjee, A. and S. Stern (2005) “Open Science: Disclosure, Research Investment and Cumulative Knowledge Production,” mimeo, Kellogg School of Management.

Murray, F. (2002) "Innovation as co-evolution of scientific and technological networks: Exploring tissue engineering," *Research Policy*, 31(8): 1389-1403.

Murray, F. and S. Stern. (2007) “Do Formal Intellectual Property Rights Hinder the Free Flow of Scientific Knowledge? An Empirical Test of the anti-Commons Hypothesis,” *Journal of Economic Behavior and Organization.*

Murray, F., P. Aghion, M. Dewatrapont, J. Kolev and S. Stern (2010). "Of Mice and Academics: The Role of Openness in Science". MIT Sloan Working Paper.

Owen-Smith, J., and W. W. Powell (2003) “The Expanding Role of University Patenting in the Life Sciences: Assessing the Importance of Experience and Connectivity,” *Research Policy*, 32(9): 1695-1711.

National Science Foundation (2010) “Science of Science and Innovation Policy (SciSIP) Additional Funding Opportunities,” accessed Nov-11-2010 from: http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=501084&org=sbe.

Rosenberg, N. (1990) “Why do firms do basic research (with their own money)?” *Research Policy*, 19(2): 165-174.

Schubert, A. and T. Braun. (1993) “Reference standards for citation based assessments,” *Scientometrics* , 26, 21-35.

Stephan, P.E. (1996) “The Economics of Science,” *Journal of Economic Literature*, 34(3): 1199-1235.

Thomson, J.A. et al. (1998). *Science* 282, 1145-1147.

Walsh, J., C. Cho and W.M. Cohen (2005) "View from the Bench: Patents, Research and Material Transfers," *Science*, September 23, 2002-2003.

Walsh, J, A. Arora and W. M. Cohen, (2003) "Effects of Research Tool Patents and Licensing on Biomedical Innovation," in W.M. Cohen and S. Merrill, eds., *Patents in the Knowledge-Based Economy*, National Academy Press.